

# Application of Area-frame sampling for agricultural statistics in China

Yaozhong Pan<sup>1</sup>, Minghua Wang<sup>2</sup>, Ge Wei<sup>2</sup>, Fenghua Wei<sup>2</sup>, Kaifen Shi<sup>2</sup>, Le Li<sup>1</sup>, Guannan Sun<sup>1</sup>

(1.College of Resources Science & Technology, State Key Laboratory of Earth Processes and Resource Ecology, Beijing Normal University,100875,Beijing,China; 2.Division of Agricultural Survey, Department of Rural Survey, National Bureau of Statistics,100826, Beijing, China)

**Abstract:** The food safety has been concerned by Chinese government all the time. The method of combining sampling survey and comprehensive survey statistics collecting is commonly used to obtain the crop area in China by the National Bureau of Statistics of China. However, the development of the rural economy in China brings new problems and challenges to this method. In this paper, the current method of crop area survey in China will be illustrated and evaluated. Based on this, a new approach will be proposed, which will be more effective and fitting. The results showed that the area-frame sampling methods will be the main trend of crop area survey in China.

**Key words:** Sampling survey; Area-frame sampling; Crop area

## 1. Introduction

Collecting the accurate crop area on time is an important assignment of Agricultural Statistic<sup>[1]</sup>. The National Bureau of Statistics of China (NBS) has gained crop planted area reliably by the survey methods of combining sampling and Stat. However, the development of rural economy brings problems and new challenges to this method. To overcome these problems, area-frame sampling method has been applied to Chinese current cropland survey.

### 1.1 The Development of Area Frame Survey in the world

The development of crop sampling survey can be categorized into three periods: 1) List frame Sampling Period: Before 1970s, List frame sampling is the dominated sampling method using the administration units or farms as sampling survey units<sup>[2]</sup>; 2) Area-frame sampling Period: From 1970 to 1990s, area-frame sampling method has been widely used in many countries, using natural parcels or regular grids as sampling units<sup>[3, 4]</sup>; 3) Area-frame sampling with remote sensing data Period: Since 2000s, the method combing the area-frame sampling and remote sensing has become a new trend because the remote sensing data can provide abundant information of land cover and land change. Many developed countries have established their area-frame sampling survey system based on remote sensing<sup>[1]</sup>.

### 1.2 The Background of Sampling Survey in China

The development of sampling in China also went through three periods. The first one is that the cropland area and crop yield collecting are in the form of symposiums in towns and villages in the early days of New China; The next one is from 1960s to 1980s, the crop area is obtained mainly by stating and submitting step by step according the administration units; And the latest one is list frame sampling according to crop area, which was combined well with the Stat. System

and improved the crop area survey in China <sup>[6]</sup>. However, with the develop of rural economy, several problems came out: 1) The descend of the cooperate of investigated famers; 2) The backward of investigate measurement and technology; 3) The heavy of task and the lack of effective supervising; 4) The quality of the statistics can't meet the increasing needs of government, enterprises and publics. Therefore, a new system need to be established, which can promote the capability of statistics service and provide smart data to meet the needs of government, enterprises and public.

NBS has been moving on. In the tenth five-year program, area-frame sampling method has been experimented in Jiangsu Province to obtain the crop area in province level. After that, the method was applied to dominated crop area estimation combining remote sensing data in another five provinces with the cooperation of Beijing Normal University (BNU). The five provinces are Hebei, Jiangsu, Anhui, Shandong, Henan; In the eleventh five-year program, Area-frame Sampling System for crop survey based on 3S technology has been established by NBS and BNU. This system has been applied to Jiangsu, Hebei, Hubei and Jilin to estimation crop area in province and county level.

### **1.3 Challenge**

However, because of the instability of agriculture ecosystem and the acquisition of remote sensing data, the Area-frame Sampling System is facing new challenges:

- 1) China is one of biggest countries of agriculture product producing, with the most complex and fragmental cultivation structure.
- 2) The acquisition of full coverage high/moderate remote sensing data in large area becomes a severe limitation for crop area survey caused by the influence of satellite revisiting cycles and weather conditions.
- 3) The lack of the mature basic data frame in China is another limitation.

It's obvious that the cropland survey can't be operational only depending on remote sensing technique. New methods based on multiple remote sensing (historical and real time data) combining remote sensing and sampling are needed for the operational cropland survey system.

## **2. The Technical Framework of Crop Area Estimation Using Remote Sensing and Statistics**

### **2.1 Concept of Technical Frame**

The content of the frame sampling is the dominated crop area in China. The crop area is estimated by multi-stage sampling methods, with the support of all kinds of agricultural census data, land cover survey data, remote sensing data and crop estimation system. The survey work was implemented by investigation teams of province and supervised by NBS. The final crop area and yield in provinces will be submitted to NBS. The concept of technical frame is as Fig.1:

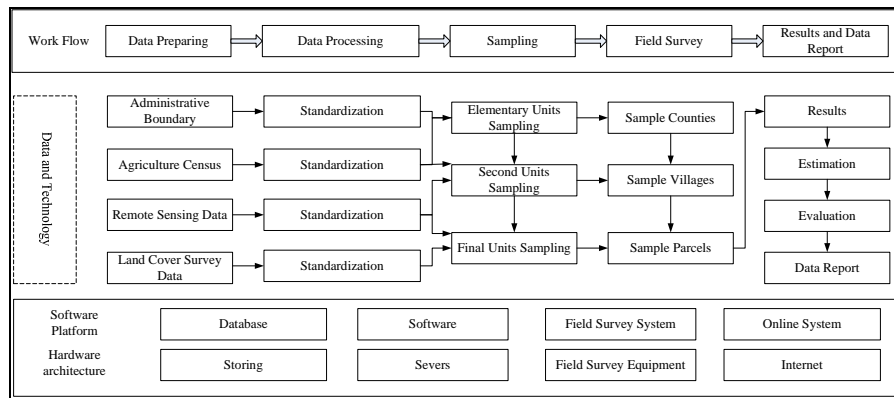


Fig. 1 The main Frame of Crop Area-Frame Sampling

The NBS implemented the crop area estimation program with area-frame sampling and issued the assignment to administrative institutions in three levels: In the nation scale: Division of Agricultural Survey of NBS is in charge of the national rural socio-economic sample survey; In the province scale: The National Bureau of Statistics Survey Corps, which is established by NBS in 31 provinces, will manage the crop sampling survey in their province separately and submitted the final results to NBS regularly; In the county level: There are 875 designated national survey counties carrying out the survey in all the sample villages.

## 2.2 Sampling Method

The sampling survey is mainly in rural households in agricultural production and management, not including state farms, such as military farms, labor camps and so on. The object is crop planted area. The multi-stage sampling method is applied to 31 provinces (autonomous regions and municipalities) separately. There are three stages: The first stage is extracting sample counties in province-wide; The second stage is collecting sample villages from sample counties; And the third stage is selecting sample parcels from sample villages. The exact processing is shown in Fig.2 as following:

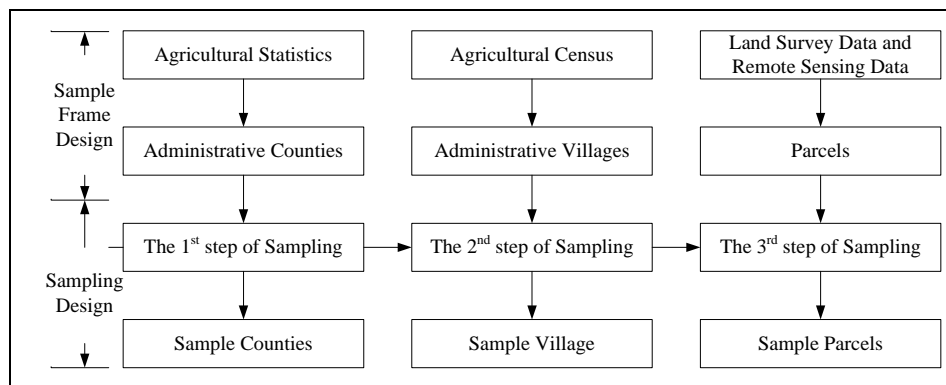


Fig.2 Sampling Design

### 2.2.1 Sampling Frame

The sampling frame designing combines the agricultural statistics and multi-source spatial information, including years of agricultural statistics, agricultural census data, the second national land survey data and multiple source remote sensing data.

The years of agricultural statistics are collected by provincial statistical departments

submitting to NBS annually; The agricultural census data is the agricultural and rural economic survey data of all Agricultural production units [6]; The second national land survey data is the production of the second national land survey by the ministry of land and resource, including land use maps in multi-scales, farmland distribution and other important production [7,8]. All of these data are very previous auxiliary data source.

In the support of the database, three stage sampling frame was designed according to the three stage sampling method. The 1<sup>st</sup> sampling frame is county sample frame with all the counties in the province as sample units. And the average crop yield of every county was calculated to be auxiliary data; The 2<sup>nd</sup> sampling frame is village sample frame with all the villages in sample counties as sample units, while the agricultural census data, land cover survey data and multiple source remote sensing data are combined to support this frame; The 3<sup>rd</sup> frame is based on parcels, which is established by the second national land cover survey. The sizes of these parcels are very different, so mergers and segmentation are needed. The Fig.3 shows the three level sampling frame of area-frame in China.

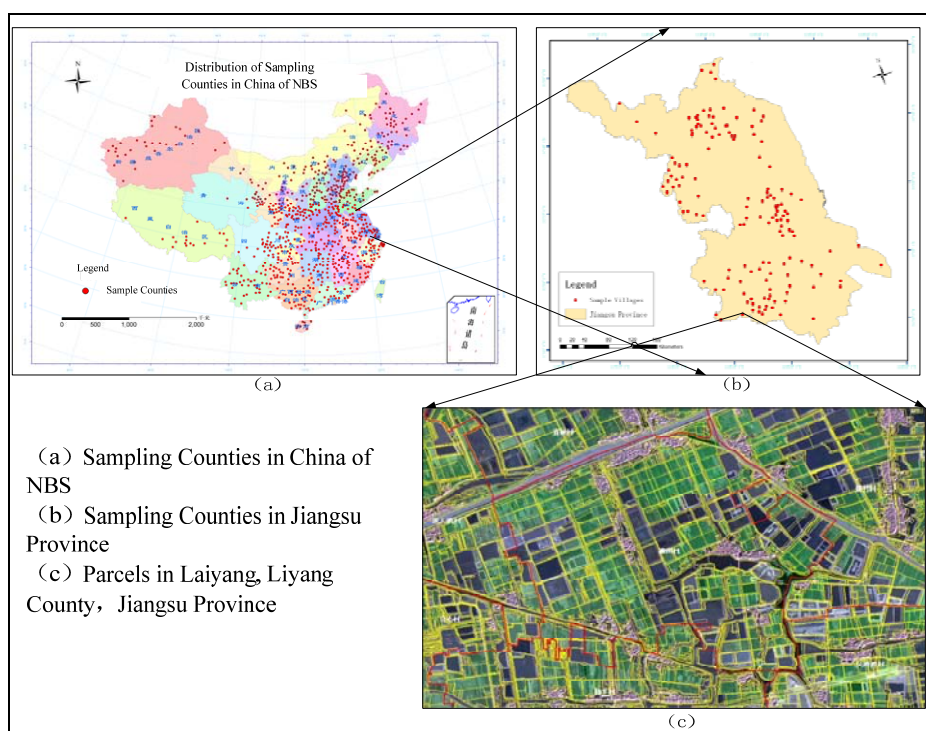


Fig.3 Sampling Frame

## 2.2.2 Sampling Design

The objects of multi-stage sampling are sampling counties, sampling villages and sampling parcels. The different objects of each stage and the relevant auxiliary data make the sampling designs different.

In the first stage, a certain population of sample counties is sampled from all the sampling provinces including autonomous regions and municipalities. The population of the sample counties accounts for 35 percent of all the counties in China. The sample counties are sampled by the symmetric interval sampling method, with the average yield per Mu in recent three years or average income of each villager as stratified auxiliary.

In the second stage, a certain population of sample villages is sampled from all the sampling

counties in each province. The population of samples is defined by the crop estimates from the agricultural census data and the estimation results from remote sensing data. The crop planted structure information in every administrative village is also very important reference. At last, the PPS sampling method is used to select sample villages, with the ratio of cultivated area as sample index.

In the third stage, a certain population of cultivated parcels is selected randomly from sample villages as final survey sample plots. The cultivate parcels are mainly defined by the second National Land Survey by the ministry of land and resource.

At last, the crop area of sample villages is estimated by survey results of sample plots, and then the crop area of sample counties is estimated by sample villages; next, to the crop area of provinces; at last, all of these are submitted to NBS.

### 3. Experiment

The experiments of area frame sampling were tested in sample counties. Two kinds of sampling method were chosen according to the difference of crop production and geographic condition: The one is simple random sampling method with standard parcels as sample units. This method is applied to Faku county in Liangning Province, Liyang county in Jiangsu Province and Puyang county in Henan Province. Every standard parcels in Faku covers around 6 hectares, and around 5 hectares in Puyang; The other one is Two-step PPS sampling method, with administrative villages (or regular grid) as the elementary sample units, the crop cultivated structure as auxiliary data and the standard parcels as the final sample units. In Fengtai county in Anhui Province, administrative villages is selected as the elementary sample units, parcels in 2 hectares as final sample units, while in Dehui county in Jilin Province, regular grids in 2000\*2000 are selected as the elementary sample units and parcels in around 7 hectares as final sample units. The site survey was carried out three times with each method. The whole area frame sampling experiments includes several processing: materials preparation, program designing, standards establishing, field survey and results organizing.

Based on the analysis of survey data, it can be found that the survey data reflected the actual cultivation and land cover change. And different sampling methods all well estimate the dominated crop area with relatively lower CV values and higher precision. However, to some small varieties crops, the CV value is less than satisfactory. The results are shown in Table.1.

Table 1 Accuracy in Sample Counties by Simple Estimation

Counties	Crop CV							
	Rice	Wheat	Maize	Other Crop	Beans	Tubers	Rape	Vegetation
Faku	31.6%	28.6%	3.4%	73.1%	32.7%	69.7%	20.7%	41.0%
Liyang	4.8%	58.4%	23.2%	43.1%	87.8%	16.8%	15.4%	20.9%
Puyang	20.5%	5.2%	56.4%	14.0%	44.1%	24.8%	19.7%	59.8%
Fengtai	6.2%	3.0%	27.7%	100.0	14.3%	59.4%	42.1%	86.1%
Dehui	43.36%	-	10.0%	-	92.5%	-	-	-

The experiments demonstrate that the processing of sampling survey can be completed and efficient. They are sample frame establishing, sample parcels surveying, parcels dividing and merging , 3S technology embedding. All of these help the survey team seize the sample parcels

fast and accurately to complete the site investigation. It is another success of the crop survey by area-frame sampling. Furthermore, the NBS will implement the sampling survey in Jiangsu and Henan province.

#### **4. Discussion**

Both of the sampling methods perform well to dominated crop, while less satisfactory to small varieties crop. It can get a higher precision by the simple random sampling method. However, the distribution of samples is scattered and the population of samples is relative large which make the surveying work heavy; The distribution of samples by grid sampling is more focused, but its accuracy can't meet the requirement of NBS; The two-stage PPS sampling combining the agricultural census can make a satisfactory accuracy with a relative fewer sample villages, which fits the organizing of NBS and can be main sampling method in the future.

In this experiment, the sample frame is defined by the parcels of the second national land survey and its standard parcels data. If the parcels are used directly, the sampling design is easier. However, the problems are: 1) The boundary of the parcels is hard to define; 2) The area of parcels is relative larger, which brings difficulty to surveying. If the standard parcels are used, the problems also came out: 1) How to define the standard parcels according to the standard area; 2) How to divide the standard parcels; 3) How to deal with the fragment parcels.

#### **5. Conclusion**

In this experiment, the area-frame sampling is demonstrated to be efficient and practical in crop area estimation. However, there are some issues need to be improved in the details and some more provinces will be chosen in the future.

#### **Reference**

- [1] Global Strategy to Improve Agricultural and Rural Statistics.
- [2] F.J.Gallego. Crop Area Estimation in the MARS Project[J]. Conference on ten years of the MARS Project, 1999,4.
- [3] R.B.MacDonald, F.G.Hall, R.B.Erba . The Use of LANDSAT Data in a Large Area Crop Inventory Experiment[R]. Symposium on Machine Processing of Remotely Sensed Data, 1975,6.
- [4] FAO. Multiple frame agricultural surveys [C]. FAO Statistical Development Series, Volume 2: Agricultural survey programs based on area frame or dual frame (area and list) sample designs, Rome, 1998.
- [5] F.J.Gallego. Using Remote Sensing for Agricultural Statistics[J]. International Statistical Review. 2005(73).
- [6]Investigation Corps of National Statistics Bureau of China. Examples of Multi-objective sampling methods [M], 2003, 9.
- [7]Ministry of Land and Resource: The State Council on the upcoming second national land survey of the notice [R]. National Land and Resources Informations. 2009(2).
- [8] Zhang Dingxiang, Liu Li, Huang Sihong, Ji Hongwei. Prospect on The 2nd National Land Survey Maps and its serving. [C]. Proceedings of 2009 Conference on Land Science.2009.